

Construindo a agenda política

Maurício Yoshida Izumi*

Universidade de São Paulo

41^o Encontro Anual da Anpocs

GT18 O presidencialismo de coalizão brasileiro

23 a 27 de Outubro de 2017

Resumo

O objetivo deste trabalho é analisar como ideias sobre políticas são construídas ao longo do tempo. Em termos empíricos, verificamos como o conteúdo presente em um projeto de lei reaparece de forma semelhante em outros projetos. Avaliamos também qual o impacto inclusão de interesses sobre a sua probabilidade de aprovação. Para isso utilizamos todos os projetos de lei apresentados durante a 54^a legislatura na Câmara dos Deputados e técnicas de análise quantitativa de textos. Mostramos que projetos que foram influenciados por outros projetos tem mais chances de aprovação, principalmente se eles forem de origem do Executivo. Sugerimos que a inclusão de interesses é o fator chave para um projeto se tornar lei.

Palavras-chave: Legislativo, Projetos de lei, Ideias políticas, Análise quantitativa de textos, Similaridade de cosseno.

*Este trabalho contou com o apoio da Fapesp, processo número 2013/24210-4.

1 Introdução

A produção de leis é uma das principais atividades de um governo. É por meio da formulação de leis que os governos são capazes de implementar as suas agendas políticas e mudar o *status quo*. Durante uma legislatura milhares de projetos de lei (PL's) são apresentados por parlamentares, por membros do Poder Executivo e do Judiciário e pela população em geral. No entanto, apenas uma pequena porcentagem desses projetos realmente se tornam uma norma jurídica. Durante a 54^a legislatura da Câmara dos Deputados (2011-2015), por exemplo, quase 7000 projetos de lei foram apresentados. Mas apenas 1% deles foram aprovados.

Embora essa taxa de aprovação seja baixa, argumentamos neste trabalho que projetos que não foram aprovados estão longe de serem irrelevantes. As ideias apresentadas neles muitas vezes influenciam o conteúdo de outros projetos. Isto é, as ideias sobre políticas presentes em um projeto apresentado por um deputado i no tempo t reaparecem em um projeto de um deputado j (ou pelo Executivo) no tempo $t + 1$.

Assim o objetivo deste trabalho é analisar como ideias sobre políticas são construídas ao longo do tempo. Em termos empíricos, verificamos como o conteúdo presente em um projeto de lei reaparece de forma semelhante em outros projetos. Avaliamos também qual o impacto da inclusão de preferências sobre a sua probabilidade de aprovação. Para isso utilizaremos todos os projetos de lei apresentados durante a 54^a legislatura na Câmara dos Deputados e técnicas de análise quantitativa de textos.

A literatura brasileira sobre produção legislativa tem se concentrado sobre o processo de tramitação individual dos projetos e os mecanismos institucionais dos presidentes. De acordo com [Figueiredo and Limongi \(1999\)](#), o Executivo domina a produção legal - mais de 70% dos projetos aprovados são de origem do Executivo - e tem uma alta taxa de sucesso na apresentação de seus projetos - mais de 85% dos projetos enviados por ele são aprovados. A explicação para esses resultados estaria nos poderes legislativos do presidente. Ao controlar a agenda, isto é, o que será votado e quando o será, os presidentes estariam em uma posição estratégica para aprovar os seus projetos.

Em uma abordagem diferente, [Freitas \(2016\)](#) mostra o papel exercido pelo Legislativo no processo de formatação das leis. Analisando os projetos apresentados pelo Executivo que foram aprovados entre 1995 e 2010, a autora mostra

que o Legislativo participa ativamente da formulação de políticas modificando, adicionando e suprimindo conteúdo dos projetos do Executivo. Os resultados apontam que, em média, 36% do texto das leis são produzidas pelo Legislativo e 64% pelo Executivo. Isto é, mesmo que o Executivo domine a agenda, isso não quer dizer que o Legislativo não participe na formatação do seu conteúdo.

[Silva and Araújo \(2010\)](#), por sua vez, mostram como o Executivo tem se apropriado de projetos do Legislativo para construir sua própria agenda. A partir de uma análise aprofundada de vinte proposições legislativas e de entrevistas realizadas com consultores legislativos, os autores dão indícios de que o Executivo se apropria do conteúdo de projetos elaborados por parlamentares ao apresentá-los por meio de medidas provisórias ou como PL's de sua autoria.

Este trabalho busca contribuir com esse debate de duas formas. A primeira contribuição é metodológica: por meio de técnicas de análise quantitativa de texto, procuramos fornecer uma medida mais objetiva da similaridade das ideias sobre políticas das proposições legislativas. A segunda contribuição é substantiva: mostramos como o fato de um projeto ter ideias que foram apresentadas em projetos anteriores aumentam as suas chances de ser aprovado. O argumento aqui é que ao apresentar um projeto muito semelhante a outro já existente, o autor não está apenas reproduzindo as suas ideias, mas sim adicionando conteúdo novo a ele. A adição de conteúdo ao projeto agrega preferências que não estavam presentes em suas versões anteriores. Assim, a inclusão de interesses seria o mecanismo associado ao aumento das chances de aprovação.

Este trabalho possui duas seções, além desta introdução e da conclusão. Na seção que se segue apresentamos os nossos dados e métodos. Para medir a similaridade das ideias sobre políticas entre os projetos de lei utilizamos técnicas de análise quantitativa de textos. Em particular, aplicamos um algoritmo frequentemente utilizado na detecção de plágio (similaridade de cosseno). Na seção 3, apresentamos os nossos resultados. Mostramos que cerca de 8% dos projetos apresentados foram influenciados por outros projetos apresentados durante a mesma legislatura. Mostramos também que projetos que foram influenciados por outros projetos tem mais chances de aprovação, principalmente se eles forem de origem do Executivo. Projetos que foram aprovados foram influenciados, em média, por 2,63 projetos. Já projetos que não foram aprovados foram influenciados, em média, por apenas 0,19 projetos.

Por fim, sugerimos que a inclusão de interesses é o fator chave para um projeto se tornar lei. Mostramos que em mais de 70% dos casos eles foram influ-

enciados por ideias formuladas por outros partidos. Isso sugere que, em geral, preferências distintas as dos autores dos projetos estão sendo incluídas. Mostramos também que em mais de dois terços das vezes os projetos adicionaram conteúdo em relação a versões anteriores. Quando comparados com aqueles projetos que suprimiram conteúdo ou que são meras cópias, os projetos que adicionaram conteúdo tem dez vezes mais chances de serem aprovados.

2 Dados e métodos

O primeiro objetivo deste trabalho é verificar como ideias sobre políticas que aparecem em um projeto de lei ressurgem em outros projetos futuros. Para isso iremos utilizar todos os projetos de lei apresentados na Câmara dos Deputados entre 1º de fevereiro de 2011 e 31 de janeiro de 2015 (54ª legislatura). Esse período coincide com o primeiro governo da presidente Dilma Rousseff (PT).

De acordo com a Constituição brasileira, um projeto de lei pode ser proposto por qualquer membro do Poder Legislativo (deputados ou senadores), por comissões (da Câmara, do Senado ou do Congresso Nacional), pelo presidente, por membros do Poder Judiciário (Supremo Tribunal Federal, Tribunais Superiores ou pelo Procurador-Geral da República) e por cidadãos comuns (por meio dos projetos de iniciativa popular). Com exceção dos projetos de lei de autoria dos senadores que iniciam a tramitação no Senado Federal, todos eles são apresentados na Câmara dos Deputados. Ao todo observamos 6732 projetos de lei apresentados na Câmara dos Deputados. Mais de 96% deles são de autoria de deputados (de forma individual ou coletiva).

Iremos trabalhar com o inteiro teor dos projetos. Isto é, o texto da lei na forma em que ele foi apresentado. Na tabela 1 temos um exemplo do PL100/2011 de autoria do deputado Walter Tosta (PMN-MG).

Em geral, os projetos de lei seguem esse formato. Primeiramente temos uma ementa onde é apresentado o objetivo do projeto. Logo em seguida temos os artigos do projeto. Em média, os projetos são constituídos por 382 palavras. Mas isso varia bastante, o desvio padrão é de 1446 palavras. O menor projeto tem 51 palavras e o maior tem mais de 85.000 palavras.

O primeiro passo para se analisar quantitativamente os textos das leis é reduzir a sua complexidade ([Grimmer and Stewart, 2013](#)). Para isso excluimos pontuações, números, palavras poucas informativas que possuem funções gra-

Tabela 1: Inteiro teor do PL100/2011

Estabelece normas para facilitar o acesso das pessoas portadoras de deficiência com com mobilidade reduzida nos centros comerciais, 'shopping centers' e hipermercados e dá outras providências.

O Congresso Nacional decreta:

Art. 1º - Os centros comerciais e 'shopping centers' e hipermercados, instalados em todo o território nacional disponibilizarão elevadores para uso exclusivo de portadores de deficiência física.

Art. 2º - Deverão ser afixadas em local de grande visibilidade, nas dependências externa e interna dos centros comerciais, 'shopping centers' e hipermercados, placas indicativas da localização dos respectivos elevadores.

Art. 3º - A não-observância desta lei sujeitará o infrator à multa pecuniária de R\$5.000,00 (cinco mil reais), valor que será corrigido anualmente Índice Geral de Preços do Mercado (IGPM), aplicada em dobro em caso de reincidência.

Art. 4º - A fiscalização do cumprimento desta lei caberá competentes do Poder Executivo.

Art. 5º - Esta lei entra em vigor na data de sua publicação.

maticais (*stopwords*) e palavras muito e pouco frequentes (que aparecem em mais de 99% e em menos de 1% dos textos). Além disso, simplificamos o vocabulário por meio de *stemming*. Isto é, reduzimos as palavras que se referem ao mesmo conceito a uma única raiz. Por exemplo, as palavras 'economia', 'economista' e 'economistas' foram reduzidas à raiz 'econom'. Com isso criamos a nossa matriz de documentos por termos (*document-term matrix*). Essa matriz é formada por 6732 linhas (número de documentos) e 1551 colunas (número de termos) cujas entradas são o número de vezes que cada termo apareceu em cada documento. Dessa forma, cada documento pode ser definido como um vetor em um espaço de dimensão 1551.

A partir desses vetores podemos criar uma medida de similaridade baseada no produto interno entre eles. Sejam \mathbf{u} e \mathbf{v} dois vetores em \mathbb{R}^{1551} , o produto interno entre \mathbf{u} e \mathbf{v} é dado por: $\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{1551} u_i v_i$. Quando os dois vetores tem valores altos nas mesmas dimensões, maior será o produto interno. Dito de outra forma, quando temos um produto interno alto, temos um indício de que a distribuição do conjunto de termos utilizada pelos documentos é similar.

Mas se definirmos o comprimento (ou norma) de um vetor \mathbf{v} como $\|\mathbf{v}\| = \sqrt{\sum_{i=1}^{1551} v_i^2}$, é fácil observar que quanto maior o comprimento dos vetores, maior será o produto interno. Quando o vetor possui valores altos em cada dimensão,

maior será o seu comprimento e, conseqüentemente, maior será o produto interno entre os vetores. Assim, documentos que utilizam palavras muito frequentes terão um comprimento maior e, como palavras frequentes tendem a ocorrer em vários documentos, maior será o produto interno entre os documentos.

A solução para esse problema é dividir o produto interno pelo produto dos comprimentos dos vetores, $\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}$. Mas sabemos da geometria analítica que $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \cos \theta$. Portanto, $\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}$. Essa medida $\cos \theta$, que nada mais é do que o cosseno do ângulo entre \mathbf{u} e \mathbf{v} , é conhecida como similaridade de cosseno (Jurafsky and Martin, 2014). Ela varia de zero a um, já que as entradas da matriz de documentos por termos assumem apenas valores positivos. Quanto mais próximo de um, maior a similaridade entre os documentos. Utilizamos essa medida para avaliar como as ideias sobre políticas de um projeto de lei ressurgem em outros projetos.

Excluimos da análise todos os projetos que tratam de homenagens e de criações de datas comemorativas por terem pouca relevância. Com isso temos um total de 6355 projetos. Comparamos cada projeto de lei com todos os outros projetos que foram apresentados anteriormente a ele. Por exemplo, se um projeto foi apresentado em 1º de abril de 2012, comparamos este projeto com todos os outros apresentados entre 1º de janeiro de 2011 e 1º de abril de 2012. Assim, podemos traçar a evolução de uma ideia ao longo da legislatura.

Seguimos uma estratégia conservadora para classificar quais projetos estão relacionados. Não consideramos simplesmente que um projeto possui uma ideia semelhante a outro projeto se o cosseno entre eles for maior do que zero. Consideramos apenas cossenos maiores do que 0,8 entre projetos que tratam do mesmo tema. Utilizamos esses critérios porque uma importante fonte de variação no uso de palavras é o estilo do documento (Lauderdale and Herzog, 2016). Textos de leis tendem a ser similares entre si pois possuem a mesma estrutura. Por se tratar de projetos de lei, eles compartilham uma similaridade que não está relacionada às ideias políticas que eles carregam. Utilizando esses critérios de classificação, estamos reduzindo as chances de classificar projetos como relacionados, quando na verdade eles não o são. Por outro, estamos deixando de classificar projetos como relacionados quando na verdade eles são relacionados. Preferimos errar pelo lado conservador.

Na tabela 2 temos uma comparação entre o PL7480/2014 de autoria do deputado Gustavo Petta (PCdoB-SP) e o PL1375/2011 da deputada Professora Dorinha Seabra Rezende (DEM-TO). A influência do texto da deputada do DEM

Tabela 2: Comparação entre o PL1375/2011 e o PL7480/2014

PL1375/2011	PL7480/2014
<p><i>Altera a redação do art. 11 da Lei nº 11.180, de 23 de setembro de 2005, no que se refere ao valor da bolsa-permanência do Programa Universidade para Todos - PROUNI.</i></p> <p>O Congresso Nacional decreta:</p> <p>Art. 1º O art. 11 da Lei nº 11.180, de 23 de setembro de 2005, passa a vigorar com a seguinte redação:</p> <p>'Art. 11. Fica autorizada a concessão de bolsa-permanência, no valor de até um salário mínimo mensal, exclusivamente para custeio das despesas educacionais, a estudante beneficiário de bolsa integral do Programa Universidade para Todos - Prouni, instituído pela Lei nº 11.096, de 13 de janeiro de 2005, matriculado em curso de turno integral, conforme critérios de concessão, distribuição, manutenção e cancelamento de bolsas a serem estabelecidos em regulamento, inclusive quanto ao aproveitamento e à frequência mínima a ser exigida do estudante.' (NR)</p> <p>Art. 2º Esta lei entra em vigor na data de sua publicação.</p>	<p><i>Altera a redação do art. 11 da Lei nº 11.180, de 23 de setembro de 2005, com relação à concessão de bolsa-permanência para estudantes beneficiários do Programa Universidade para Todos (Prouni).</i></p> <p>O Congresso Nacional decreta:</p> <p>Art. 1º O art. 11 da Lei nº 11.180, de 23 de setembro de 2005, passa a vigorar com a seguinte redação:</p> <p>'Art. 11. Fica autorizada a concessão de bolsa-permanência, até o valor equivalente ao praticado na política federal de concessão de bolsas de iniciação científica, exclusivamente para custeio das despesas educacionais, a estudantes beneficiários de bolsa integral do Programa Universidade para Todos (Prouni), instituído pela Lei nº 11.096, de 13 de janeiro de 2005, conforme critérios de concessão, distribuição, manutenção e cancelamento de bolsas a serem estabelecidos em regulamento, inclusive quanto ao aproveitamento e à frequência mínima a ser exigida do estudante.</p> <p>Parágrafo único. Os critérios de concessão referidos no caput considerarão especialmente a situação de impossibilidade de compatibilidade entre a frequência ao curso, em turno parcial ou integral, e o exercício de atividade remunerada, no caso de o estudante não contar com renda própria ou familiar suficiente para prover sua subsistência.' (NR)</p> <p>Art. 2º Esta lei entra em vigor na data de sua publicação.</p>

(PL1375/2011) sobre o projeto do deputado do PCdoB (PL7480/2014) é evidente. A ementa dos projetos possui o mesmo conteúdo com pequenas mudanças de redação. O 1º e o 2º artigos são idênticos. Apenas o artigo 11 possui mudanças de conteúdo. Além disso, o projeto de 2014 acrescentou conteúdo novo (Parágrafo único) em relação ao projeto de 2011. Neste exemplo, o cosseno do ângulo entre os dois vetores é de 0,85. Na seção seguinte apresentamos os nossos resultados.

3 Resultados

A literatura brasileira não tem dado atenção para o conteúdo textual dos projetos e a sua relação com a probabilidade de aprovação. Para o caso norte-americano, alguns trabalhos mostraram o poder preditivo dos textos das leis na probabilidade de sobrevivência nas comissões (Yano et al., 2012), na previsão do comportamento em votações nominais (Gerrish and Blei, 2011) e na probabilidade de aprovação (Nay, 2016).

Outros trabalhos estudaram a difusão de ideias na produção legislativa das assembleias estaduais. Garrett and Jansa (2015) e Kroeger (2015) mostraram o papel fundamental dos grupos de interesse. De acordo com os autores, legisladores estaduais frequentemente recorrem à *expertise* desses grupos no processo de formulação de políticas. Burgess et al. (2016) mostraram que legisladores estaduais aumentam a sua produtividade utilizando projetos formulados em outros estados e por grupos de interesse.

Por meio da análise dos projetos de lei, diversos trabalhos traçaram o fluxo de ideias no interior congresso (Wilkerson et al., 2015; Li et al., 2014; Smith et al., 2014). Os autores mostraram como muitas ideias propostas em um projeto acabam por se tornar parte de outros projetos. Um bom exemplo disso é o *Obamacare*, cuja origem é um plano de reforma do sistema de saúde proposta pelos Republicanos em 1993 (Wilkerson et al., 2015).

Para o caso brasileiro, verificamos que 92% dos projetos apresentados na 54ª legislatura da Câmara dos Deputados não foram influenciados por nenhum outro projeto apresentado durante o mesmo período. No entanto, é importante ressaltar que isso não quer dizer que esses projetos sejam totalmente novos. Existe a possibilidade deles terem sido influenciados por projetos apresentados em legislaturas passadas.

Tabela 3: Resultados da regressão logística

Variável	Coefficiente	Erro padrão
Intercepto	-5,12	0.16
Número de projetos similares	0,21	0.03
Executivo	4.93	0.40

Nota: unidade de análise = projeto; variável dependente = número de projetos similares; número de observações = 6355.

Em média, os projetos foram influenciados por 0,20 projetos com um desvio padrão de 1,17. O projeto com o maior número de conexões é o PL7944/2014 de autoria do deputado Alceu Moreira (PMDB-RS) com 24 projetos relacionados. Entre os projetos aprovados essa média sobe para 1,46 com um desvio padrão de 2,63 (o mínimo é 0 e o máximo é 8). Entre aqueles que não foram aprovados a média é menor, 0,19 projetos relacionados com um desvio padrão de 1,14. Mas a amplitude é maior, o mínimo é 0 e o máximo é 24. Essa simples comparação de médias sugere que projetos que foram influenciados por um maior número de projetos tendem a ter mais chances de serem aprovados.

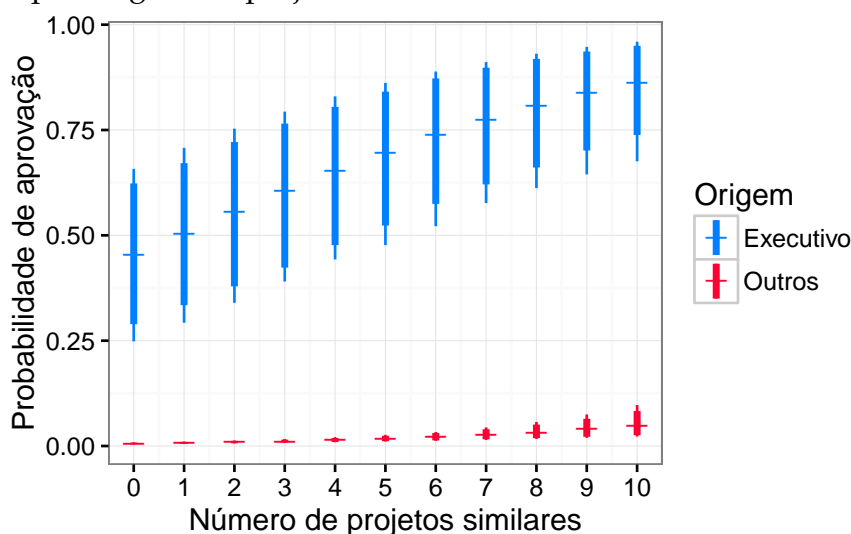
Para testarmos essa ideia de forma mais acurada propomos um modelo de regressão logística tendo como unidade de análise o projeto. A nossa variável dependente é um indicador que vale 1, caso o projeto tenha sido aprovado no período analisado, e 0, caso contrário. A principal variável explicativa é o número de projetos que o influenciou. Esperamos que quanto maior o número de projetos similares, maiores as chances de aprovação.

Utilizamos como variável de controle um indicador de se o projeto é de origem do Executivo (1, se o projeto é de origem do Executivo; 0, caso contrário). Como sabemos, projetos do Executivo tem mais chances de serem aprovados devido aos poderes de agenda dos presidentes (Figueiredo and Limongi, 1999). Assim, esperamos um impacto positivo dessa variável sobre a probabilidade de aprovação do projeto.

Na tabela 3 apresentamos os resultados. Como podemos observar, todas as variáveis tem o sinal esperado e são estatisticamente significantes aos níveis convencionais. Do ponto de vista substantivo, podemos dizer que projetos do Executivo tem mais chances de serem aprovados. E mais importante, projetos que foram influenciados por outros projetos com ideias semelhantes também possuem mais chances de se tornarem lei.

Para termos uma ideia mais clara desse impacto, no gráfico 1 temos a rela-

Figura 1: Impacto do número de projetos similares sobre a probabilidade de aprovação por origem do projeto



Nota: Os valores preditos foram calculados utilizando o pacote Zelig do R (Imai et al., 2008). As linhas na horizontal indicam os valores preditos. As linhas grossas na vertical indicam o intervalo de confiança de 90% e as linhas finas na vertical indicam o intervalo de confiança de 95%.

ção entre a probabilidade de aprovação e o número de projetos similares, controlando pela origem do projeto. Quando um projeto do Executivo não possui nenhum outro projeto similar, ele tem 45% de chances de ser aprovado (IC95% : [0,25;0,66]). Sob as mesmas condições, um projeto cuja origem não é o Executivo, essa probabilidade é bem menor, 0,6% (IC95% : [0,004;0,008]). Esse resultado era esperado já que os poderes legislativos dos presidentes influenciam diretamente no andamento das proposições.

O simples fato de um projeto do Executivo ter apenas um outro projeto similar aumenta a probabilidade de aprovação para 50% (IC95% : [0,29;0,71]). Se forem dois projetos similares, temos 56% de chances de aprovação (IC95% : [0,34;0,75]). Mas o mesmo não ocorre com os projetos cuja origem não é o Executivo. Um projeto que foi influenciado por outros 10 projetos similares tem apenas 5% de chances de ser aprovado (IC95% : [0,02;0,10]).

Sugerimos que ao apresentar um projeto muito semelhante a outro já existente, o autor não está apenas reproduzindo as suas ideias, mas adicionando conteúdo novo a ele. Ao trazer novas ideias a um projeto já existente, o autor estaria agregando preferências em torno dele. Assim, o mecanismo que estaria associado ao aumento das chances de aprovação seria a inclusão de interesses.

Do ponto de vista empírico, isso tem duas implicações. A primeira é a de que os projetos devem ser influenciados por projetos apresentados por indivíduos com preferências distintas. Um projeto de um deputado X que influencia um projeto de um deputado Y só estará agregando interesses se o deputado X tiver preferências diferentes das do deputado Y. Se as preferências forem iguais, nenhum interesse estará sendo agregado. A segunda implicação é a de que os projetos que foram influenciados adicionam conteúdo em relação aqueles que os influenciaram. Se conteúdo novo não é adicionado, isto é, se os projetos são meras cópias de versões anteriores, não temos a inclusão de interesses.

Vamos avaliar a primeira implicação. Se os projetos forem influenciados por projetos de indivíduos que possuem as mesmas preferências, não podemos dizer que houve a inclusão de interesses. Por exemplo, se um deputado apresenta um projeto que foi influenciado por projetos apresentados por ele mesmo no passado, não houve inclusão de ideias novas. A agregação de preferências ocorre quando ideias de deputados que possuem posições políticas diferentes são adicionadas ao projeto.

Uma forma de verificar isso é olhar para o partido dos autores dos projetos. É razoável imaginar que indivíduos de diferentes partidos tenham interesses distintos. Dos 534 projetos que foram influenciados por pelo menos um projeto, 506 são de autoria de deputados. Nesse subconjunto somos capazes de saber qual o partido do autor do projeto e quais são os partidos dos autores dos projetos que os influenciaram. Assim, podemos verificar se os projetos são influenciados por projetos do próprio partido ou se são influenciados por projetos de outros partidos.

Desses 506 projetos, 59% foram influenciados somente por projetos cuja autoria não é o partido do próprio parlamentar; 28% foram influenciados somente por projetos cuja autoria é um parlamentar do próprio partido; e 13% foram influenciados por projetos tanto do próprio partido quanto de partidos diferentes. Em outras palavras, em mais de 70% dos casos, os projetos que foram influenciados por outros projetos levaram em consideração as ideias de outros partidos. Isso sugere que as ideias de parlamentares com preferências distintas são levadas em consideração na formatação das leis.

Vamos avaliar a segunda implicação. Embora seja difícil verificar se um projeto traz ideias novas em relação a outros projetos, podemos ter uma ideia olhando para o comprimento dos textos. Se os projetos forem simples cópias de versões anteriores, esperamos que o comprimento dos projetos seja igual. Por

outro lado, a adição de novas ideias levaria a um aumento do texto em termos do número de palavras. Podemos então calcular o número de palavras em cada projeto de lei e o número médio de palavras nos projetos que influenciaram cada projeto. Por exemplo, o PL1337/2011 tem 175 palavras e foi influenciado pelo PL581/2011 (106 palavras) e pelo PL877/2011 (165 palavras). Portanto, a média do número de palavras dos projetos que influenciaram o PL1337/2011 é igual a 135,5. Consequentemente, o PL1337/2011 adicionou conteúdo com relação aos projetos anteriores.

O tamanho médio dos projetos que foram influenciados por pelo menos um projeto é de 517 palavras (com um desvio padrão de 1615). Já o tamanho médio dos projetos que influenciaram os projetos é de 369 (com um desvio padrão de 1086). A comparação entre essas médias sugere que os projetos que foram influenciados por outros projetos são maiores do que aqueles que os influenciaram. O teste de diferenças de médias unilateral à direita indica que essa diferença não é fruto do acaso (valor de $p = 0,039$). Em termos substantivos, em média, conteúdo novo foi adicionado às versões anteriores.

Podemos classificar os 534 projetos que foram influenciados por pelo menos um projeto em três categorias. Um projeto pode ser menor do que aqueles que o influenciou. Neste caso, houve supressão de conteúdo da versão anterior para a versão mais recente. Um projeto também pode ser maior do que aqueles que o influenciou. Isso ocorre quando uma versão é aprimorada com a adição de conteúdo. Por fim, um projeto pode ser igual àqueles que o influenciou. Neste caso, o projeto é uma simples cópia. Utilizamos o seguinte critério para classificar os projetos:

$$\text{Projeto}_i = \begin{cases} \text{menor,} & \text{se } \|\text{Projeto}_i\| + 10 < \|\overline{\text{Similares}}_i\| \\ \text{igual,} & \text{se } \|\text{Projeto}_i\| + 10 \geq \|\overline{\text{Similares}}_i\| \geq \|\text{Projeto}_i\| - 10 \\ \text{maior,} & \text{se } \|\text{Projeto}_i\| - 10 > \|\overline{\text{Similares}}_i\| \end{cases}$$

onde $\|\text{Projeto}_i\|$ é o comprimento do projeto $i = \{1, \dots, 534\}$ e $\|\overline{\text{Similares}}_i\|$ é a média do comprimento dos projetos que influenciaram i . Consideramos um projeto igual àqueles que o influenciou sempre que os seus comprimentos forem iguais, adicionando uma margem de 10 palavras. Essa margem tem o intuito de garantir que diferenças muito pequenas nos textos não sejam classificadas como uma adição ou supressão de conteúdo.

Seguindo essa classificação temos que 62% dos projetos tem um comprimento maior do que os projetos que os influenciaram. Em outras palavras, em quase dois terços das vezes, os projetos que foram influenciados por outros projetos adicionaram conteúdo em relação as versões anteriores. Em 17% dos casos houve supressão de conteúdo. Isto é, o comprimento do projeto é menor do que daqueles que o influenciou. E em 21% das vezes os projetos mantiveram o mesmo tamanho.

Calculando a razão de chances entre projetos que adicionaram conteúdo e projetos que reduziram ou que permaneceram com o mesmo tamanho encontramos um valor de 9,67. Isto quer dizer que projetos que são maiores do que aqueles que os influenciaram tem quase 10 vezes mais chances de serem aprovados do que aqueles que suprimiram conteúdo ou que são simples cópias. Em termos substantivos, isso indica que projetos que trazem novas ideias em relação a versões anteriores tem mais chances de se tornarem lei do que projetos que são cópias ou que suprimiram conteúdo.

4 Conclusões

O objetivo deste trabalho foi analisar como ideias sobre políticas são construídas ao longo do tempo. Em termos empíricos, verificamos como o conteúdo presente em um projeto de lei reaparece de forma semelhante em outros projetos. Avaliamos também qual o impacto da inclusão de interesses nos projetos de lei sobre a sua probabilidade de aprovação. Para isso utilizamos todos os projetos de lei apresentados durante a 54^a legislatura na Câmara dos Deputados e técnicas de análise quantitativa de textos.

Mostramos que cerca de 8% dos projetos apresentados foram influenciados por outros projetos apresentados durante a mesma legislatura. Mostramos também que projetos que foram influenciados por outros projetos tem mais chances de aprovação, principalmente se eles forem de origem do Executivo. Sugerimos que a inclusão de diferentes interesses é o fator chave para um projeto se tornar lei.

O argumento aqui é que ao apresentar um projeto muito semelhante a outro já existente, o autor não está apenas reproduzindo as suas ideias, mas sim adicionando conteúdo novo a ele. Dessa forma, a agregação de preferências em torno do projeto é o que aumentaria as suas chances de aprovação. Dentre os

projetos que foram influenciados por pelo menos um projeto, mais de 70% foram influenciados por projetos de outros partido. Além disso, quase dois terços deles adicionaram conteúdo novo à uma versão anterior. Mostramos, por fim, que projetos que foram agregaram interesses tem dez vezes mais chances de serem aprovados do que aqueles que suprimiram conteúdo ou que são meras cópias de outros projetos.

Os estudos legislativos no Brasil tem focado as suas análises sobre o processo de tramitação individual dos projetos e sobre os mecanismos institucionais dos presidentes. Essas análises assumem que os projetos são independentes uns dos outros. Assim, dois projetos que compartilham o mesmo conteúdo teriam as mesmas chances de aprovação, mesmo que um agregue mais interesses do que o outro. Este trabalho buscou uma abordagem diferente. Procuramos verificar como ideias políticas evoluem ao longo do tempo. Ou seja, como ideias que surgiram em um projeto em um determinado momento são incluídas em projetos futuros. Nesse sentido, entendemos a história de projetos políticos como a construção de ideias ao longo do processo legislativo.

Referências

- Burgess, M., Giraudy, E., Katz-Samuels, J., Walsh, J., Willis, D., Haynes, L., and Ghani, R. (2016). The legislative influence detector: Finding text reuse in state legislation. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 57–66. ACM.
- Figueiredo, A. and Limongi, F. (1999). *Executivo e Legislativo na nova ordem constitucional*. Editora FGV.
- Freitas, A. (2016). *O Presidencialismo da Coalizão*. Fundação Konrad Adenauer.
- Garrett, K. N. and Jansa, J. M. (2015). Interest group influence in policy diffusion networks. *State Politics & Policy Quarterly*, 15(3):387–417.
- Gerrish, S. and Blei, D. M. (2011). Predicting legislative roll calls from text. In *Proceedings of the 28th international conference on machine learning (icml-11)*, pages 489–496.
- Grimmer, J. and Stewart, B. M. (2013). Text as data: The promise and pitfalls of

- automatic content analysis methods for political texts. *Political analysis*, pages 267–297.
- Imai, K., King, G., and Lau, O. (2008). Toward a common framework for statistical analysis and development. *Journal of Computational Graphics and Statistics*, 17(4):892–913.
- Jurafsky, D. and Martin, J. H. (2014). *Speech and language processing*, volume 3. Pearson.
- Kroeger, M. (2015). Plagiarizing policy: Model legislation in state legislatures. In *Working Paper*.
- Lauderdale, B. and Herzog, A. (2016). Measuring political positions from legislative speech. *Political Analysis*, 24(3):374–394.
- Li, W. P., Larochelle, D., and Lo, A. W. (2014). Estimating policy trajectories during the financial crisis. In *Working Paper*.
- Nay, J. J. (2016). Predicting and understanding law-making with machine learning. In *Working Paper*.
- Silva, R. and Araújo, S. (2010). Apropriação da agenda do legislativo: como aferir esse fenômeno. *Texto para Discussão. Centro de Estudos da Consultoria, Senado Federal*, 76.
- Smith, D. A., Cordell, R., Dillon, E. M., Stramp, N., and Wilkerson, J. (2014). Detecting and modeling local text reuse. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*, pages 183–192. IEEE Press.
- Wilkerson, J., Smith, D., and Stramp, N. (2015). Tracing the flow of policy ideas in legislatures: A text reuse approach. *American Journal of Political Science*, 59(4):943–956.
- Yano, T., Smith, N. A., and Wilkerson, J. D. (2012). Textual predictors of bill survival in congressional committees. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 793–802. Association for Computational Linguistics.